Pacific Northwest
NATIONAL LABORATORY

PNNL-30511

# De Novo Design of Molecular Recognition for Sequence Defined Polymers

September 2020

Marcel D. Baer

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

## DISCLAIMER



PACIFIC NORTHWEST NATIONAL LABORATORY
*operated by*
BATTELLE
*for the*
UNITED STATES DEPARTMENT OF ENERGY
*under Contract DE-AC05-76RL01830*

**Printed in the United States of America**

**Available to DOE and DOE contractors from the Office of Scientific and Technical Information,**
**P.O. Box 62, Oak Ridge, TN 37831-0062**
www.osti.gov
**ph: (865) 576-8401**
**fax: (865) 576-5728**
**email: reports@osti.gov**

**Available to the public from the National Technical Information Service**
**5301 Shawnee Rd., Alexandria, VA 22312**
**ph: (800) 553-NTIS (6847)**
**or (703) 605-6000**
**email: info@ntis.gov**
**Online ordering: http://www.ntis.gov**

PNNL-30511

# De Novo Design of Molecular Recognition for Sequence Defined Polymers

September 2020

Marcel D. Baer

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

## Overview

The proposed research aims at developing a computational framework that allows the de novo design of sequence defined polymers for molecular recognition. This framework will include the prediction of the 1) accessible backbone conformational space in solution environments for small to medium polymer length and 2) how side chains of different size and functionality can be used to design convergent or divergent receptor configurations. A structural database inspired by the protein database will be build that can be used to construct larger scaffolds based on smaller sequences.

## Introduction

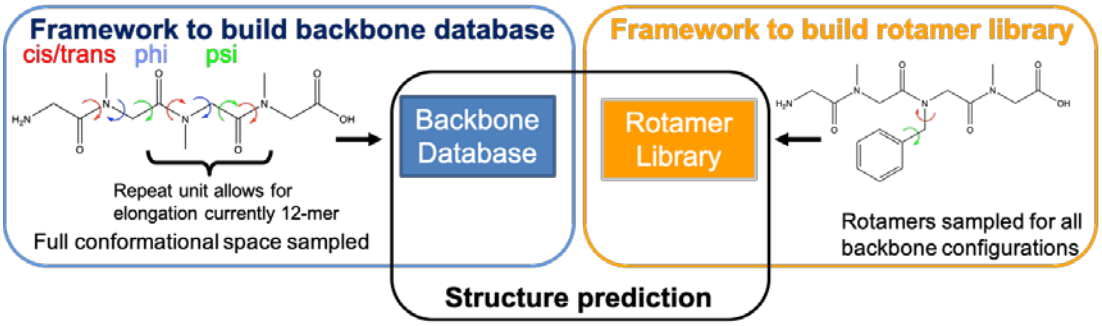
Synthetic receptors are employed for in many applications like separations, imaging and sensing, catalysis, pharmaceutical activity and functional materials. These applications play crucial roles within the frontier technology areas of health care, environmental remediation, nanotechnology, and advanced materials. Each application requires a synthetic receptor that is endowed with a specific set of functional properties. The functions are triggered by a molecular recognition event that involves association of the receptor with the target molecule. The associated complex is held together by a collection of weak and reversible bonds. Typically, the association alters the chemical and physical properties of the binding partners enabling the subsequent functions to occur.

In recent years, advances in biological, chemical, physical and computational methodologies have provided researchers with the tools not only to identify and characterize interacting molecules, but also to understand the general rules of molecular recognition. In drug design, computer driven design methods are widely used, and range from virtual screening of known compounds to de novo design without prior knowledge of previously synthesized active ligands. While the problem of designing synthetic receptors for molecules is essentially the inverse of drug design, the computational design of synthetic receptors is underdeveloped. Here we address this challenge to develop a computational capability for de novo design for molecular recognition using sequence defined polymers (peptiods, peptides and triazine based polymers) utilizing different design principles.

### Results

A computational framework for the prediction of the accessible conformational space of sequence defined polymer was developed for peptoids, using our recent developed force fields here at PNNL. A conformational database for a minimal backbone system for up to 12 units was developed based on all atom molecular dynamics simulations using parallel tempering in explicit solvent. For this chain length, the full conformational space can be explored. As the amount of data is immense, cluster algorithms are used to define stable local conformations which then are stored in the structure database. This database can be search for cavities that would allow the binding of substrates. Subsequentially, sidechains around these cavities can then be modified to introduce stabilizing interactions. The introduction of new side chains requires a new parameterization and also a local relaxation. This approach turned out to be to computational and time intensive.

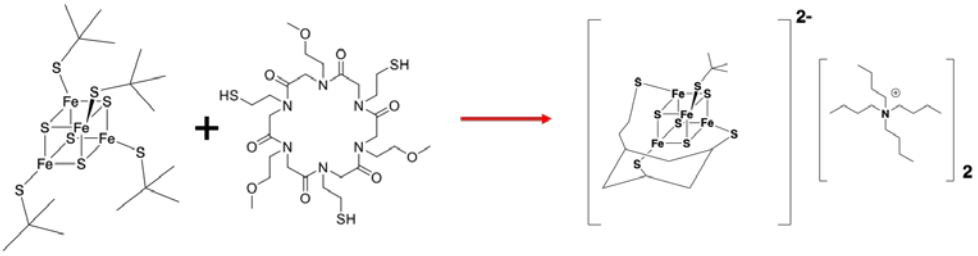
We recently showed that the backbone conformational space for peptoids can be reliable described on the semi-empirical level allowing an efficient and accurate sampling without the need of force field parameterizations for the unusual side chains. This is mainly due to the fact that unlike peptides no stabilizing hydrogen bonding exit for the backbone, which is usually not sufficiently described with semi-empirical methods. We developed a fully automated framework to generate backbone conformations that allows not only the use of peptoids, but other sequence defined polymers as well.




**Scheme:** Structure prediction framework based of sampling of backbone conformational space and automated generation of rotamer libraries for side chains. Fully automated using semi empirical methods to avoid the need of force field parameterization.

The structure of the monomer is provided as SMILES and the connectivity of the backbone needs to be specified. The SMILES string is converted into xyz coordinates using obabel and hydrogens are added. The conformations are generated using developed tcl-scripts within VMD. The same framework can also be used to sample possible sidechain conformations of the side chains. This allows for the generation of a backbone database and rotamer library that can be used for structure prediction.

To complement this framework, a protocol was developed to calculate rotameric libraries that can be directly used in the ROSETTA program suit. This enables enhanced sampling by swapping different functionalities of single side chain using well established monte carlo algorithms. Even though straight forward for peptoids, due to the similarity to peptides this is an easy extension, this is not directly applicable for other sequence defined polymers.



**Scheme:** Predicted cyclic peptoid structure that can bind and stabilize $Fe_4S_4$ in solution synthesis route.

As a non-trivial test case, a prediction for a peptoid that can bind and stabilize a $Fe_4S_4$ was made to be tested experimentally. The structure and predicted conformation in which the sulfur atoms are positioned to bind to three corners of the FeS-cluster is shown in Scheme 2. As a first step to validate the prediction a synthesis scheme was developed to yield $Fe_4S_4$ stabilized by S-t-butyl. This precursor can be used to perform ligand exchange reactions with peptoids. Ligand exchange of S-t-butyl ligand with thiol containing peptoids was achieved and confirmed by NMR. The formed $Fe_4S_4$-peptoid complex is stable and shows changes in the redox-properties of the cluster as confirmed by cyclic voltammetry. The predicted peptoid was synthesized, but the ligand exchange reaction was not successful yet.

# Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

***www.pnnl.gov***